



**USER GUIDE:  
HCUP SUPPLEMENTAL FILES FOR REVISIT ANALYSES**

**January 11, 2010**

## Table of Contents

|   |    |
|---|----|
| Introduction .....  | 1  |
| Development of the HCUP Revisit Files .....   | 2  |
| Verified Patient Identifiers (visitLink) .....  | 2  |
| Calculating the Days to Event (daysToEvent) .....   | 3  |
| HCUP Revisit Files Data Elements .....  | 4  |
| Relationship to HCUP State Databases .....  | 4  |
| Availability of Verified Patient Identifiers Varies By State .....  | 5  |
| Selecting States for a Revisit/Readmission Analysis .....   | 6  |
| Using the HCUP Revisit Files .....  | 7  |
| Usage Examples .....  | 7  |
| Usage Example #1: Assigning Patient Characteristics .....   | 7  |
| Usage Example #2: Revisits for Selected Patients .....  | 7  |
| Usage Example #3: Preceding Visits to any Hospital Setting for Selected Patients .....  | 8  |
| Cautionary Note: Transfers and Possible Duplicates .....  | 9  |
| Appendix A: Available HCUP Supplemental Files for Revisit Analyses .....  | 11 |
| Appendix B: Example of Assignment of visitLink and daysToEvent .....  | 12 |
| Appendix C: Consistency of Person Identifiers across Years .....  | 13 |
| Appendix D: Consistency of Verified Revisit Information .....   | 14 |
| Percentage of Discharges with a Verified Person Number (visitLink), All Available HCUP<br>Databases within a State, 2006-2007 ..... | 15 |
| Percentage of Discharges with a Verified Person Number (visitLink), All Available HCUP<br>Databases within a State, 2005-2006 ..... | 16 |
| Percentage of Discharges with a Verified Person Number (visitLink), All Available HCUP<br>Databases within a State, 2004-2005 ..... | 17 |
| Percentage of Discharges with a Verified Person Number (visitLink), All Available HCUP<br>Databases within a State, 2003-2004 ..... | 18 |
| Appendix E: SAS Code For Usage Examples .....   | 19 |
| Usage Example #1: Assigning Patient Characteristics .....   | 19 |
| Usage Example #2: Revisits for Selected Patients .....  | 21 |
| Usage Example #3: Preceding Visits to any Hospital Setting for Selected Patients .....  | 25 |

## INTRODUCTION

The AHRQ Healthcare Cost and Utilization Project (HCUP) databases capture information on inpatient (IP), emergency department (ED) and ambulatory surgery (AS) encounters in U.S. community hospitals. These databases are often characterized as being “discharge-level” files, meaning that each record in a database represents one discharge abstract from a hospital setting, which can be an IP, ED, or AS visit. Thus, if the same individual visited the hospital multiple times in a given year, the HCUP databases would include separate records in the respective HCUP database for each visit. Many times researchers may be interested in knowing how many visits a distinct patient had rather than simply the number of overall hospital visits. Studying multiple visits is becoming increasingly common as hospital readmission rates are seen as an important indicator of the quality of medical care. To facilitate analyses that focus on multiple hospital stays by the same person, AHRQ created a set of supplemental files that can be linked to the HCUP state-level databases to track multiple (repeat) patient visits in the hospital setting while adhering to strict privacy regulations. This user guide discusses the methodology used to develop these supplemental files and describes how to apply the files to the HCUP databases. Appendices provide detail on which states, data types and years are available, in addition to verification statistics. Several SAS coding excerpts are also provided to facilitate the use of these files.

We refer to these files as the “HCUP Supplemental Files for Revisit Analyses” (herein referred to as the “Revisit Files”). These supplemental files were labeled “revisit” files rather than “readmission” files intentionally. The term “revisits” implies multiple health care encounters for a particular patient that are not limited solely to IP stays. Most health care research on hospital *readmissions* has focused solely on the IP setting – i.e., tracking multiple hospital admissions in the IP setting by the same person. The HCUP Revisit Files expand on traditional readmission analyses by allowing researchers to study multiple patient visits to the hospital, regardless of whether the visit resulted in a hospital readmission. In other words, these supplemental files enable more than identification of hospital readmissions – they also enable tracking of patients admitted to the hospital following an ambulatory surgery (AS) or an emergency department (ED) visit and patients who made multiple trips to the ED. Note that revisits may occur for any reason (i.e., they may not be related) and can be separated by days or years. The determining factor in classifying health care events as revisits is that they represent services for the same individual.

In contrast, readmissions are sequential hospital admissions for a related reason, and usually within a specified time frame. Studying readmissions can be difficult as researchers must understand whether patients are admitted for expected follow-up treatment, or, conversely, for unexpected complications. In addition, multiple hospital visits for the same patient may, in fact, be unrelated - and, therefore, not considered a “readmission.” Identifying readmissions requires specific criteria for the inclusion of events, such as type of condition and appropriate elapsed time. For example, a study of readmissions for congestive heart failure (CHF) may require that the principal reason for the hospitalization, ED visit, or ambulatory surgery is related to CHF and may also require that the time elapsed between events is not longer than predetermined number of days. The HCUP Revisit Files contain key data elements, such as the days between multiple visits, that can assist analysts in making informed decisions about whether repeat patient visits qualify as readmissions.

The HCUP Revisit Files, used in combination with HCUP state-level databases, enable analysts to link hospital visits that belong to a unique person, determine the elapsed time between visits, and evaluate valuable clinical information on the HCUP discharge abstract .

These files afford analysts the flexibility of performing patient-level analyses within and across hospital settings and time periods, without compromising patient confidentiality. Finally, these files allow the analyst to determine their definition of a readmission or revisit for the purposes of their study. This User Guide documents the creation of these files and provides guidance on how to best utilize the files in revisit analyses.

## **HCUP Databases**

HCUP develops and maintains a family of health care databases, related software tools, products and support services. HCUP features the largest collection of multi-year hospital care data in the United States, containing a wealth of all-payer, encounter-level information beginning in 1988. AHRQ relies on vital partnerships among Federal, State, and Industry associations to produce HCUP resources. HCUP databases integrate the data collected by state governments, hospital associations, private data organizations, and the Federal government to create a national health care information resource of IP, AS, and ED data.

The HCUP Revisit Files are designed to be used exclusively with the HCUP state-level databases:

- The State Inpatient Databases (SID), which contain inpatient discharge records from non-Federal hospitals in participating states
- The State Emergency Department Databases (SEDD), which contain data from emergency department encounters from hospital-affiliated emergency departments
- The State Ambulatory Surgery Databases (SASD), which include data from ambulatory care encounters from hospital-affiliated and sometimes freestanding ambulatory surgery sites.

The Revisit Files are unique within state and data year. Users will need to merge the data elements on the HCUP Revisit Files to the corresponding SID, SASD, or SEDD for any analysis (further described in the section on Using HCUP Revisit Files).

Note: HCUP Revisit Files are available for some, but not all, state-level databases – SID, SASD, and SEDD – starting in calendar year 2003. Appendix A lists the availability of the supplemental files by state, year, and data type.

## **DEVELOPMENT OF THE HCUP REVISIT FILES**

Development of the HCUP Revisit Files required state-level HCUP files to contain a unique encrypted patient identifier which enable tracking of unique patients within and across years. Only some of the HCUP statewide data organizations provide this information to HCUP. Each state employs a distinct methodology in producing their encrypted person numbers.

### **Verified Patient Identifiers (visitLink)**

An instrumental part of constructing the HCUP Revisit Files was verifying that the encrypted person numbers accurately represented a unique person in the HCUP state databases. As part of the verification process, the patient's date of birth and gender were used to qualify the encrypted patient numbers and uniquely identify a person. A new verified person number (**visitLink**) was assigned for each unique combination of the qualifying information (encrypted person number, date of birth, and gender). Consider the following example: Five records have

the same encrypted person numbers, but two records have one date of birth and gender, and the remaining three records have a different, but consistent, date of birth and gender. The two records with identical identifying information have one value of **visitLink**, and the other three records have a different value of **visitLink**. Appendix B contains examples of the assignment of **visitLink** for different scenarios.

No verified person number is assigned if any of the three pieces of information was missing (i.e., **visitLink** is missing). Additionally, no verified person number is assigned if there were more than 40 hospital visits in a given calendar year with the same qualifying information. This second qualification excluded less than 0.5 percent of the person numbers and aimed to eliminate person numbers used for multiple people. Appendix B contains examples of the assignment of **visitLink** for different scenarios.

While the term “verified person number” is used to describe the information in the HCUP data element **visitLink**, the values are not recognizable as specific patient information. **VisitLink** does not include the values of the encrypted person number, date of birth, or gender.

The **visitLink** variable is created each year for all discharges across all available databases for the particular state. For example, if a state provides SID, SASD, as well as SEDD, the **visitLink** is verified for all discharges across all three databases for that data year and previous years of data, as appropriate.

### Calculating the Days to Event (**daysToEvent**)

For a verified person number (i.e., non-missing **visitLink**) with more than one hospital visit, the elapsed days between visits are calculated as the difference between the two visit dates. This information is often useful for determining readmissions for a specific condition (i.e., 30-day readmissions, 7-day readmissions, etc.) While this information is critical for defining readmissions, the use of admission and discharge dates is highly restricted per Health Insurance Portability and Accountability Act (HIPPA) guidelines.

To comply with HIPPA guidelines and ensure patient confidentiality, no “date” information is released on the HCUP Revisit Files. A timing variable (**daysToEvent**) was calculated consistently for each verified person number (**visitLink**) based on a randomly assigned “start date.” Each verified person number is assigned a unique start date that is used to calculate **daysToEvent** for all visits associated with that **visitLink** value. The variable **daysToEvent** is the difference between the visit’s admission date and the start date associated with the **visitLink**.

The calculation of days between visits is the difference of **daysToEvent** between two selected visits for a unique verified person number (**visitLink**). For example, consider a patient with congestive heart failure that has a hospital admission on 1/10/2008 and an ED visit on 1/25/2008. If the **daysToEvent** value is “9” for the 1/10/08 admission and the **daysToEvent** value is “24” for the 1/25/2008 ED visit, then the number of days between the two visits is 15 days ( $24 - 9 = 15$ ).

The lowest value of **daysToEvent** will be on the first or earliest event for a patient. It is important to remember that if patient A has a value of 605 for **daysToEvent** and patient B has a value of 300 for **daysToEvent**, patient B’s event did not necessarily take place prior to patient A’s event – in fact, Patient B’s **daysToEvent** value has no relation to Patient A’s **daysToEvent** value. Because of the use of a random start date in the calculation of **daysToEvent**, the value

of **daysToEvent** cannot be compared across patients. Appendix B contains examples of the assignment of **daysToEvent** for different scenarios.

## HCUP REVISIT FILES DATA ELEMENTS

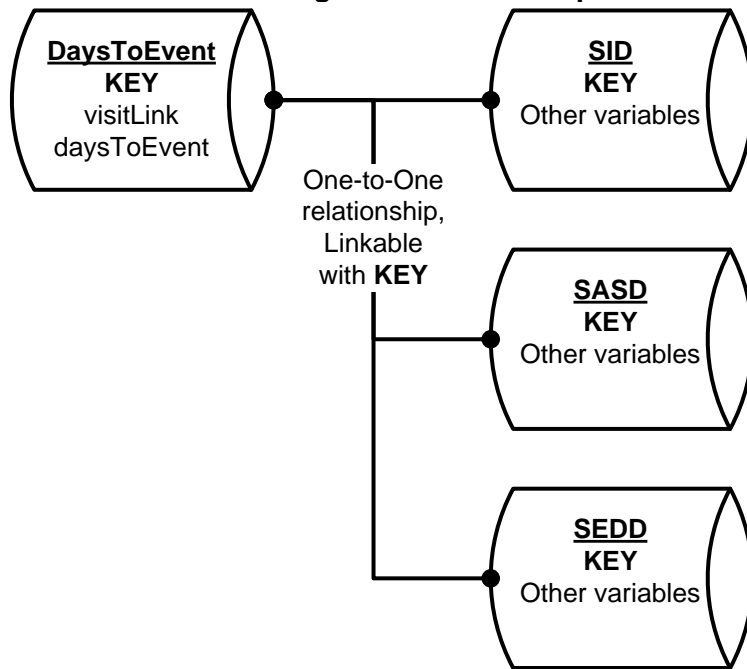
The HCUP Revisit Files include only three data elements:

- **KEY**: unique HCUP record identifier for each state, year, and database that is assigned during state database processing for linking HCUP data files
- **visitLink**: linkage variable for all events associated with a unique patient that is assigned during construction of the supplemental revisit files and based on a unique combination of person numbers, date of birth, and gender
- **daysToEvent**: the number of days from a randomly chosen "start date" to the admission date for a specific health care visit for an individual. The start date is randomly assigned for each unique patient. As a result, **daysToEvent** will be consistently calculated for all of a patient's linked events, regardless of year (i.e., all visits with the same value of **visitLink**). The daysToEvent variable is assigned during construction of the supplemental revisit files. This information can be linked to the HCUP SID, SASD, and/or SEDD files based on the KEY variable.

## RELATIONSHIP TO HCUP STATE DATABASES

The HCUP Revisit Files are state and year specific and linkable to the corresponding state's available HCUP state data (SID, SASD, and SEDD) for that year (Figure 1). For example, for Nebraska there is one HCUP Revisit File for the 2006 data year. Researchers can add the **visitLink** and **daysToEvent** data elements to the 2006 Nebraska SID, SASD, and/or SEDD files by linking on the **KEY** data element.

**Figure 1. Relationship Between HCUP Databases**



## AVAILABILITY OF VERIFIED PATIENT IDENTIFIERS VARIES BY STATE

The availability of verified person numbers for specific populations and settings varies by state and should be considered prior to any analysis. Table 1 shows the range of the percentage of verified revisit information across 15 states in 2005-2006 for selected patient characteristics, expected payer, and hospital characteristics.

**Table 1. Range of Percentage of Records with Verified Revisit Information**

|                                   |                      | All Events |       |        |        |
|-----------------------------------|----------------------|------------|-------|--------|--------|
|                                   |                      | Min        | Q1    | Median | Max    |
| <b>Overall</b>                    |                      | 65.4%      | 84.7% | 90.5%  | 100.0% |
| <b>By Patient Characteristics</b> |                      |            |       |        |        |
| <b>Age Group</b>                  | 0                    | 11.0%      | 26.8% | 49.1%  | 100.0% |
|                                   | 1-17                 | 35.1%      | 54.1% | 73.3%  | 100.0% |
|                                   | 18-44                | 69.5%      | 92.6% | 95.9%  | 100.0% |
|                                   | 45-64                | 69.2%      | 96.4% | 97.5%  | 100.0% |
|                                   | 65+                  | 69.8%      | 97.2% | 98.8%  | 100.0% |
| <b>Gender</b>                     | Male                 | 64.3%      | 82.6% | 88.9%  | 100.0% |
|                                   | Female               | 66.2%      | 86.6% | 91.8%  | 100.0% |
| <b>Patient Income</b>             | Quartile 1 (lowest)  | 59.3%      | 87.2% | 91.5%  | 100.0% |
|                                   | Quartile 2           | 66.9%      | 86.7% | 91.5%  | 100.0% |
|                                   | Quartile 3           | 70.1%      | 86.1% | 90.9%  | 100.0% |
|                                   | Quartile 4 (highest) | 70.5%      | 83.4% | 88.9%  | 100.0% |
| <b>Expected Payer</b>             | Medicare             | 69.8%      | 97.2% | 99.1%  | 100.0% |
|                                   | Medicaid             | 55.4%      | 71.7% | 85.9%  | 100.0% |
|                                   | Private insurance    | 66.5%      | 82.2% | 89.4%  | 100.0% |

|                                    |                            | All Events |       |        |        |
|------------------------------------|----------------------------|------------|-------|--------|--------|
|                                    |                            | Min        | Q1    | Median | Max    |
|                                    | Self-pay                   | 66.9%      | 79.8% | 88.2%  | 100.0% |
|                                    | No Charge                  | 54.2%      | 79.0% | 86.5%  | 100.0% |
|                                    | Other                      | 44.0%      | 87.6% | 93.7%  | 100.0% |
| <b>By Hospital Characteristics</b> |                            |            |       |        |        |
| <b>Hospital Ownership</b>          | Government, nonfederal     | 72.5%      | 81.7% | 90.5%  | 100.0% |
|                                    | Private, not-profit        | 67.9%      | 84.7% | 90.0%  | 100.0% |
|                                    | Private, invest-own        | 57.5%      | 86.0% | 92.5%  | 100.0% |
| <b>Hospital Location</b>           | Large central metropolitan | 78.0%      | 82.4% | 87.8%  | 100.0% |
|                                    | Large fringe metropolitan  | 36.3%      | 83.1% | 89.2%  | 100.0% |
|                                    | Medium metropolitan        | 74.3%      | 85.4% | 92.0%  | 100.0% |
|                                    | Small metropolitan         | 45.6%      | 87.8% | 94.4%  | 100.0% |
|                                    | Micropolitan               | 69.9%      | 88.7% | 95.4%  | 100.0% |
|                                    | Noncore                    | 35.1%      | 83.4% | 94.8%  | 100.0% |
| <b>Hospital Bed Size</b>           | <100                       | 56.7%      | 87.1% | 92.5%  | 100.0% |
|                                    | 100-299                    | 68.7%      | 84.3% | 91.1%  | 100.0% |
|                                    | 300-499                    | 70.1%      | 89.7% | 92.2%  | 100.0% |
|                                    | 500+                       | 72.7%      | 84.3% | 92.5%  | 100.0% |

Source: HCUP State Inpatient Databases, 15 States, 2005-2006

In most cases, verification rates across patient and hospital characteristics and across selected diagnosis and procedure categories were consistent with the overall verification rates. For example, the first and second quartiles of the verification percentage overall were 84.7 percent and 90.5 percent, respectively. The first and second quartiles of the verification percentage for patients from hospitals in large fringe metropolitan areas were 83.1 percent and 89.2 percent, respectively.

Some notable exceptions include:

- Newborns (age 0) – The median of the verification rates across the 15 states was only 49.1 percent.
- Children and adolescents (age 1-17) – The first quartile for verification rates was 54.1 percent and the median was 73.3 percent.
- Expected payer of Medicaid and No Charge – The first quartile for both was less than 80 percent and the median was about 86 percent.

Revisit/readmission analyses for pediatric conditions and certain payers may only be appropriate in selected states.

## SELECTING STATES FOR A REVISIT/READMISSION ANALYSIS

When selecting which states to use for a revisit analysis, please reference the following resources:

- Appendix A provides the list of all states, years, and databases with HCUP Revisit Files.
- Appendix C lists states that have inconsistent coding across data years of the source person numbers provided by the HCUP Partners and should not be used for analyses that span certain years.



- Appendix D provides verification rates by state and year that should be used to determine which HCUP states are best for specific types of revisit or readmission analysis.

## USING THE HCUP REVISIT FILES

Using the HCUP Revisit Files involves four basic steps:

1. For a given state and year, merge the HCUP Revisit File with the corresponding SID, SASD, or SEDD by the data element **KEY** to add the revisit data elements **visitLink** and **daysToEvent**.
2. Select patients of interest.
3. Use **visitLink** to identify all events for a patient. The same unique value of **visitLink** is coded on all records for an individual patient. Records with missing values for the **visitLink** variable will be a mixture of patients with unknown revisit information. It may be appropriate to exclude these records from the analysis.
4. Use **daysToEvent** to sequentially order the visits for a patient and to calculate the time between two visits for a patient. If the **daysToEvent** is 5 on one event and 35 on another, the time between events is 30 (35-5).

### Usage Examples

Use of the HCUP Revisit Files is relatively straightforward. Below are three examples of applying these files to research topics.

#### *Usage Example #1: Assigning Patient Characteristics*

Researchers may want to group patients by specific patient characteristics, such as a patient's age or insurance status. When a patient's health care experience includes more than one hospital event, categorizing the patient may be problematic. This difficulty arises because some patient characteristics may change over time. To assign attributes based on when a person began receiving services, consider the following steps:

- Data should first be grouped by patient, in service date order (specifically, the HCUP state-level data file merged with the HCUP Revisit File are sorted by **visitLink** and **daysToEvent**).
- All records for a patient are then sequentially examined in order to select the first valid, non-missing value for each patient characteristic (age, gender, race, income quartile, location, expected payer).
- The selected attribute(s) are then applied to all events for the patient.

A SAS coding example of how attributes can be assigned is shown in Appendix E.

#### *Usage Example #2: Revisits for Selected Patients*

This example counts the number of related events for selected patients with a specific diagnosis and calculates a number of statistics, including days between the initial event and the first subsequent event, by setting. This example focuses on revisits for diabetes, but can easily be adapted to any diagnoses.

The example looks for a “clean period,” measured in months, with no hospital events for an individual patient for the specified condition. Use of a “clean period” for counting readmissions is optional. Sometimes when identifying an episode of care, rather than straight utilization, a period of time during which the patient has not been admitted or treated is required. The first event after the “clean period” is considered the index event. Any event in a predetermined period of time after the index event becomes part of the “episode.”

For illustrative purposes, we selected “diabetes mellitus with complications” (CCS diagnosis category 50) as the condition and required a clean period of 6 months. The steps are:

1. Combine the event and revisit data
  - a. Limit data to linkable patients (a non-missing **visitLink** available)
  - b. Select all events with the specified condition (diabetes)
2. Sort the combined events into patient (**visitLink**) and service sequence (**daysToEvent**) order
3. Find individuals with two or more events
  - a. Find patients with a clean period before their first diabetes event
    - i. On the first event for a patient, the service must be after clean period, defined as the first six months of the data year
    - ii. If the first event was prior to the clean period month, look for a clean period on subsequent events by testing the number of days between the current event and the preceding event
  - b. When a clean period is identified
    - i. Count the number of events after the clean period
    - ii. Determine the settings of the first and second events, and calculate the number of days between the first and second event
  - c. Summarize the processing counts
4. Calculate statistics (distribution) for the number of patient events
5. Summarize revisits by the initial and second service settings.

SAS programming code for this example is found in Appendix E.

### ***Usage Example #3: Preceding Visits to any Hospital Setting for Selected Patients***

This example identifies patients' hospital events that precede CABG surgery (CCS procedure category 44), regardless of service setting, and summarizes counts by principal diagnosis. The steps are:

1. Combine the event and revisit data
  - a. Limit data to linkable patients (a non-missing **visitLink** available)
  - b. Identify events with the specified procedure (CABG)
2. Sort the combined events into patient (**visitLink**) and service sequence (**daysToEvent**) order
3. For patients who received CABG surgery, select all events prior to the surgery

#### 4. Summarize prior events by primary diagnosis and setting

SAS programming code for this example is found in Appendix E.

#### Cautionary Note: Transfers and Possible Duplicates

The HCUP Revisit Files allow an analyst to identify which records in the SID, SASD, and SEDD belong to the same person, as well as the time between events for that person. An analyst still must decide how to handle the following two types of scenarios:

- Transfers – when a patient is transferred from one acute care hospital to another
- Duplicates – when a record for the same event occurs twice in the HCUP file.

In the SID, there will be two different records if a patient is transferred from one hospital to another. The following can be used to identify the two SID records:

- Same person (**visitLink** is the same on two records)
- Disposition indicating transferred out (**DISPuniform** = 2)
- Admission source indicating transfer in (**ASOURCE** = 2)
- Discharge date of one record is the same as the admission date of another (**daysToEvent** plus the length of stay of the first record equals the daysToEvent of the second record)
- Different hospital (**DSHOSPID** is different).

Analysts conducting patient-level analyses need to decide how best to use the above information to identify transfers. The coding of admission source and discharge disposition is not always consistent with the timing of events identified by **daysToEvent** (i.e., **daysToEvent** may identify two records as two parts of a transfer, but either the disposition or admission source is not coded as such). Table 2 demonstrates the range in the percentage of discharges identified as transfers using different schemes.

**Table 2. Range of Percentage of Inpatient Discharges Identified as Transfers**

| <b>Scheme to identify transfers</b>   | <b>Minimum Value Across 15 States</b> | <b>Maximum Value Across 15 States</b> |
|---|---------------------------------------|---------------------------------------|
| Percentage of records identified as transfers using one source of information:        |                                       |                                       |
| Dates   | 1.80%                                 | 6.41%                                 |
| Discharged as a transfer to another acute care hospital (DISPuniform=2)               | 1.36%                                 | 3.11%                                 |
| Admitted as a transfer from an acute care hospital (ASOURCE=2)                        | 0.85%                                 | 5.30%                                 |
| Percentage of records identified as transfers using two sources of information:       |                                       |                                       |
| Dates and DISPuniform=2   | 0.70%                                 | 2.07%                                 |
| Percentage of records identified as transfers using all three sources of information: |                                       |                                       |
| Dates, DISPuniform=2, ASOURCE=2   | 0.22%                                 | 1.50%                                 |

Source: HCUP State Inpatient Databases, 15 States, 2006

For some analyses it may be best to combine the two records from a transfer into one by summing the lengths of stay and total charges.

The HCUP SID, SASD, and SEDD occasionally have multiple records for the same person (**visitLink**) with the same **daysToEvent** and length of stay (**LOS**). These duplicate records may or may not have the similar charge and diagnostic information. Analysts should decide how best to handle such records.

In addition, HCUP made an explicit decision to duplicate records across the SEDD and SASD when a record indicated that the patient received services in both settings. In this case, the analyst will need to decide how to include and account for these cases. The effect of these duplicated records varies by state from less than 1% in California SASD to about 15% in Tennessee SASD.

## APPENDIX A: AVAILABLE HCUP SUPPLEMENTAL FILES FOR REVISIT ANALYSES

HCUP Revisit Files are available from the HCUP Central Distributor for the following states, years, and databases.

| State and Data Type   | Data Year |      |      |      |      |
|-----------------------|-----------|------|------|------|------|
|                       | 2003      | 2004 | 2005 | 2006 | 2007 |
| <b>Arizona</b>        |           |      |      |      |      |
| SID                   | Yes       | Yes  | Yes  | Yes  | Yes  |
| SEDD                  | --        | --   | Yes  | Yes  | Yes  |
| <b>Arkansas</b>       |           |      |      |      |      |
| SID                   | --        | Yes  | Yes  | Yes  | Yes  |
| <b>California</b>     |           |      |      |      |      |
| SID                   | Yes       | Yes  | Yes  | Yes  | Yes  |
| SASD                  | --        | --   | Yes  | Yes  | Yes  |
| SEDD                  | --        | --   | Yes  | Yes  | Yes  |
| <b>Florida</b>        |           |      |      |      |      |
| SID                   | --        | Yes  | Yes  | Yes  | Yes  |
| SASD                  | --        | Yes  | Yes  | Yes  | Yes  |
| SEDD                  | --        | --   | Yes  | Yes  | Yes  |
| <b>Hawaii</b>         |           |      |      |      |      |
| SID                   | --        | --   | --   | Yes  | Yes  |
| SEDD                  | --        | --   | --   | Yes  | Yes  |
| <b>Nebraska</b>       |           |      |      |      |      |
| SID                   | Yes       | Yes  | Yes  | Yes  | Yes  |
| SASD                  | Yes       | Yes  | Yes  | Yes  | Yes  |
| SEDD                  | Yes       | Yes  | Yes  | Yes  | Yes  |
| <b>Nevada</b>         |           |      |      |      |      |
| SID                   | Yes       | Yes  | Yes  | Yes  | Yes  |
| <b>North Carolina</b> |           |      |      |      |      |
| SID                   | Yes       | Yes  | Yes  | Yes  | Yes  |
| SASD                  | Yes       | Yes  | Yes  | Yes  | Yes  |
| <b>New York</b>       |           |      |      |      |      |
| SID                   | Yes       | Yes  | Yes  | Yes  | Yes  |
| SASD                  | Yes       | Yes  | Yes  | Yes  | Yes  |
| <b>South Carolina</b> |           |      |      |      |      |
| SID                   | Yes       | Yes  | Yes  | Yes  | Yes  |
| SASD                  | Yes       | Yes  | Yes  | Yes  | Yes  |
| SEDD                  | Yes       | Yes  | Yes  | Yes  | Yes  |
| <b>Utah</b>           |           |      |      |      |      |
| SID                   | Yes       | Yes  | Yes  | Yes  | Yes  |
| SASD                  | Yes       | Yes  | Yes  | Yes  | Yes  |
| SEDD                  | Yes       | Yes  | Yes  | Yes  | Yes  |
| <b>Washington</b>     |           |      |      |      |      |
| SID                   | Yes       | Yes  | Yes  | Yes  | Yes  |

## APPENDIX B: EXAMPLE OF ASSIGNMENT OF VISITLINK AND DAYSTOEVENT

The following table lists examples of the assignment of **visitLink** and **daysToEvent** in different scenarios.

| Observation | Example PNUM | Example DOB   | GENDER | Example visitLink | Calculation for daysToEvent   | Notes  |
|-------------|--------------|---------------|--------|-------------------|---|--|
| 1           | A            | 15-Jan-1960   | M      | 11111             | Calculated using same randomly selected start date for observations 1 and 2     | Same PNUM, DOB, gender for observations 1 and 2 - same vistLink  |
| 2           | A            | 15-Jan-1960   | M      | 11111             |   |  |
| 3           | A            | 15-Jan-1960   | F      | 11112             | Calculated using different randomly selected start date                         | Same PNUM and DOB as observations 1 and 2 , but different gender - new vistLink                                  |
| 4           | B            | 1-May-1940    | F      | 11113             | Calculated using same randomly selected start date for observations 4 through 6 | Same PNUM, DOB, gender for observations 4 through 6 - same vistLink  |
| 5           | B            | 1-May-1940    | F      | 11113             |   |  |
| 6           | B            | 1-May-1940    | F      | 11113             |   |  |
| 7           | B            | 15-Jun-1945   | F      | 11114             | Calculated using same randomly selected start date for observations 7 through 9 | Same PNUM and gender as observations 4 through 6 , but different DOB - new vistLink                              |
| 8           | B            | 15-Jun-1945   | F      | 11114             |   |  |
| 9           | B            | 15-Jun-1945   | F      | 11114             |   |  |
| 10          | C            | 1-Dec-1980    | M      | 11115             | Calculated using same randomly selected start date for observations 10 and 11   | Same PNUM, DOB, gender for observations 10 and 11 - same vistLink  |
| 11          | C            | 1-Dec-1980    | M      | 11115             |   |  |
| 12          | C            | Not Available | M      | Missing           | Missing   | Same PNUM and gender as observations 10 and 11, but missing date of birth - vistLink and DaysToEvent are missing |

## APPENDIX C: CONSISTENCY OF PERSON IDENTIFIERS ACROSS YEARS

The HCUP data element **visitLink** is based on the encrypted person numbers supplied by the HCUP Partners. Between 2003 and 2007, the following Partners changed the coding scheme between data years. This means that the **visitLink** created for these states can only be used to track patients within years that had similar coding of the encrypted person number.

- **Florida** – Coding of person numbers changed in 2005.
  - Use **visitLink** to track patients within 2004.
  - Do not use visitLink to track patients before and after 2005.
- **New York** – Coding of person numbers changed in 2003 and then again in 2004.
  - Use **visitLink** to track patients within 2003.
  - Use **visitLink** to track patients within 2004.
  - Use **visitLink** to track patients from 2005 forward.
  - Do not use visitLink to track patients between 2003, 2004, and later years.
- **Washington** – Coding of person numbers changed in 2007.
  - Use **visitLink** to track patients from 2003-2006.
  - Use **visitLink** to track patients within 2007.
  - Do not use visitLink to track patients before and after 2007.

## APPENDIX D: CONSISTENCY OF VERIFIED REVISIT INFORMATION

The consistency of the verified person numbers is evaluated when the HCUP Revisit Files are created for a state. Two years of data and all available data types are used for the evaluation. The year-specific tables on the following pages detail the percentage of discharges with a verified person number (**visitLink**) for the 2003-2007 HCUP Revisits Files. If the percentage of verified patient identifiers within a patient or hospital characteristic is less than 80 percent, then the percentage is listed in red. The tables allow the analyst to determine the best possible states for a revisit analysis.

Researchers should use the HCUP Revisit Files with caution when looking at revisits for specific patient populations that do not have a high percentage of verified person numbers. If studying pediatric conditions, consider states with a high percentage of verified person identifiers for ages under 18. If a proposed study was specific to payers, select states with a high percentage of verified person identifiers by payer.



**Percentage of Discharges with a Verified Person Number (visitLink), All Available HCUP Databases within a State, 2006-2007**

| Type of Discharge              | AZ    | AR    | CA    | FL    | HI    | NC    | NE     | NV    | NY    | SC    | UT    | WA    |
|--------------------------------|-------|-------|-------|-------|-------|-------|--------|-------|-------|-------|-------|-------|
| <b>Overall:</b>                | 96.6% | 96.2% | 78.6% | 90.0% | 96.6% | 72.3% | 99.7%  | 80.8% | 95.0% | 99.1% | 81.7% | 67.4% |
| <b>Age Group:</b>              |       |       |       |       |       |       |        |       |       |       |       |       |
| 0                              | 96.7% | 81.4% | 16.8% | 30.4% | 96.7% | 21.0% | 100.0% | 41.3% | 86.5% | 99.2% | 43.2% | 70.3% |
| 1-17                           | 94.1% | 85.3% | 49.1% | 70.9% | 94.1% | 57.1% | 100.0% | 47.1% | 87.7% | 99.9% | 39.9% | 71.0% |
| 18-44                          | 95.8% | 97.0% | 84.0% | 94.2% | 95.8% | 73.2% | 99.4%  | 84.8% | 95.8% | 98.7% | 92.1% | 67.9% |
| 45-64                          | 98.0% | 99.4% | 91.8% | 97.3% | 98.0% | 69.3% | 99.5%  | 91.1% | 96.6% | 99.0% | 95.7% | 64.9% |
| 65+                            | 99.3% | 99.8% | 95.5% | 98.7% | 99.3% | 70.0% | 99.8%  | 93.2% | 99.4% | 99.3% | 97.4% | 66.8% |
| <b>Expected Primary Payer:</b> |       |       |       |       |       |       |        |       |       |       |       |       |
| Medicare                       | 99.5% | 99.8% | 96.5% | 99.2% | 99.5% | 71.3% | 99.6%  | 92.5% | 98.6% | 98.8% | 97.9% | 66.7% |
| Medicaid                       | 98.9% | 89.4% | 65.7% | 80.7% | 98.9% | 63.4% | 99.6%  | 63.1% | 90.2% | 98.6% | 65.9% | 68.9% |
| Private insurance              | 99.5% | 96.3% | 77.1% | 89.1% | 99.5% | 66.0% | 99.8%  | 82.9% | 95.9% | 99.5% | 82.2% | 67.2% |
| Self-pay                       | 79.4% | 93.9% | 72.2% | 85.9% | 79.4% | 71.7% | 99.6%  | 64.6% | 91.7% | 99.1% | 78.5% | 67.7% |
| No charge                      | 59.1% | 97.5% | NA    | 85.9% | 59.1% | NA    | NA     | 50.4% | 77.1% | NA    | 81.4% | 65.3% |
| Other                          | 96.0% | 95.9% | 81.7% | 91.5% | 96.0% | 54.4% | 99.8%  | 79.7% | 97.5% | 99.3% | 84.4% | 67.0% |
| <b>Hospital Location:</b>      |       |       |       |       |       |       |        |       |       |       |       |       |
| Large central metropolitan     | 95.9% | NA    | 77.7% | 88.1% | 95.9% | 84.1% | NA     | 78.6% | 92.9% | NA    | 81.6% | 67.3% |
| Large fringe metropolitan      | 96.9% | 99.3% | 79.3% | 88.6% | 96.9% | 53.9% | NA     | NA    | 96.6% | 99.8% | 57.8% | 69.5% |
| Medium metropolitan            | 97.3% | 95.5% | 79.9% | 90.9% | 97.3% | 72.6% | 99.7%  | 87.6% | 97.2% | 99.3% | 83.7% | 65.7% |
| Small metropolitan             | 98.9% | 97.9% | 81.8% | 92.8% | 98.9% | 44.6% | NA     | NA    | 97.8% | 98.7% | 84.4% | 66.2% |
| Micropolitan                   | 97.6% | 97.0% | 88.0% | 94.5% | 97.6% | 74.3% | 99.7%  | 89.7% | 98.4% | 98.8% | 79.7% | 67.0% |
| Noncore                        | 96.9% | 95.6% | 87.1% | 95.4% | 96.9% | 48.3% | 99.5%  | 84.6% | 97.2% | 98.7% | 78.0% | 66.8% |
| <b>Hospital Bed Size:</b>      |       |       |       |       |       |       |        |       |       |       |       |       |
| <100                           | 97.7% | 96.1% | 81.0% | 92.3% | 97.7% | 61.9% | 99.7%  | 90.8% | 98.0% | 98.3% | 79.9% | 67.5% |
| 100-299                        | 96.9% | 95.6% | 77.7% | 90.4% | 96.9% | 72.3% | 99.7%  | 85.9% | 96.2% | 99.2% | 81.0% | 67.1% |
| 300-499                        | 97.9% | 96.7% | 79.6% | 91.5% | 97.9% | 67.7% | 99.6%  | 91.3% | 94.3% | 99.2% | 87.7% | 68.9% |
| 500+                           | 94.6% | 98.0% | 80.2% | 86.4% | 94.6% | 71.6% | 99.7%  | 69.2% | 93.7% | 99.2% | 88.2% | 65.2% |

**Percentage of Discharges with a Verified Person Number (visitLink), All Available HCUP Databases within a State, 2005-2006**

| Type of Discharge              | AZ    | AR    | CA    | FL    | NC    | NE     | NV    | NY    | SC     | UT    | WA     |
|--------------------------------|-------|-------|-------|-------|-------|--------|-------|-------|--------|-------|--------|
| <b>Overall:</b>                | 96.2% | 96.2% | 79.0% | 90.5% | 65.4% | 99.5%  | 88.0% | 94.9% | 99.5%  | 81.8% | 100.0% |
| <b>Age Group:</b>              |       |       |       |       |       |        |       |       |        |       |        |
| 0                              | 94.3% | 81.0% | 17.4% | 33.9% | 17.6% | 100.0% | 49.1% | 86.3% | 99.8%  | 43.3% | 100.0% |
| 1-17                           | 92.4% | 84.7% | 50.8% | 73.3% | 54.9% | 100.0% | 59.0% | 87.5% | 100.0% | 41.1% | 100.0% |
| 18-44                          | 95.9% | 97.0% | 84.1% | 94.4% | 69.5% | 99.2%  | 92.7% | 95.8% | 99.1%  | 92.5% | 100.0% |
| 45-64                          | 98.2% | 99.4% | 92.0% | 97.5% | 69.2% | 99.2%  | 97.6% | 96.6% | 99.4%  | 95.8% | 100.0% |
| 65+                            | 99.3% | 99.7% | 95.9% | 98.8% | 69.8% | 99.7%  | 98.5% | 99.3% | 99.8%  | 96.8% | 100.0% |
| <b>Expected Primary Payer:</b> |       |       |       |       |       |        |       |       |        |       |        |
| Medicare                       | 99.5% | 99.8% | 96.7% | 99.2% | 69.8% | 99.5%  | 99.1% | 98.6% | 99.4%  | 96.9% | 100.0% |
| Medicaid                       | 97.7% | 89.2% | 67.2% | 83.2% | 55.4% | 99.6%  | 74.8% | 90.3% | 99.0%  | 65.3% | 100.0% |
| Private insurance              | 98.8% | 96.0% | 77.2% | 89.4% | 66.5% | 99.5%  | 87.9% | 95.8% | 99.8%  | 82.5% | 100.0% |
| Self-pay                       | 80.3% | 93.9% | 71.9% | 86.0% | 66.9% | 99.8%  | 78.6% | 91.7% | 99.4%  | 79.3% | 100.0% |
| No charge                      | 77.1% | 92.2% | NA    | 86.0% | NA    | NA     | 69.7% | 82.0% | NA     | 70.9% | 100.0% |
| Other                          | 95.5% | 96.6% | 82.5% | 92.4% | 54.2% | 99.6%  | 86.3% | 97.6% | 99.7%  | 85.3% | 100.0% |
| <b>Hospital Location:</b>      |       |       |       |       |       |        |       |       |        |       |        |
| Large central metropolitan     | 96.1% | NA    | 78.0% | 88.7% | 86.4% | NA     | 87.9% | 92.9% | NA     | 82.4% | 100.0% |
| Large fringe metropolitan      | 94.6% | 99.0% | 79.2% | 89.2% | 50.2% | NA     | NA    | 96.4% | 99.8%  | 36.3% | 100.0% |
| Medium metropolitan            | 96.3% | 95.7% | 80.4% | 91.3% | 74.3% | 99.5%  | 88.3% | 97.2% | 99.5%  | 83.6% | 100.0% |
| Small metropolitan             | 97.2% | 97.9% | 82.9% | 93.2% | 45.6% | NA     | NA    | 97.7% | 99.2%  | 85.0% | 100.0% |
| Micropolitan                   | 96.6% | 96.5% | 88.5% | 94.8% | 69.9% | 99.7%  | 89.2% | 98.4% | 99.7%  | 80.0% | 100.0% |
| Noncore                        | 94.8% | 95.5% | 87.6% | 95.9% | 35.1% | 99.5%  | 86.8% | 97.1% | 99.1%  | 75.7% | 100.0% |
| <b>Hospital Bed Size:</b>      |       |       |       |       |       |        |       |       |        |       |        |
| <100                           | 96.6% | 96.1% | 81.5% | 92.3% | 56.7% | 99.6%  | 92.5% | 98.1% | 99.3%  | 78.6% | 100.0% |
| 100-299                        | 95.9% | 95.4% | 77.6% | 91.1% | 68.7% | 99.7%  | 90.6% | 96.0% | 99.5%  | 81.4% | 100.0% |
| 300-499                        | 97.6% | 97.2% | 80.7% | 91.9% | 70.1% | 98.9%  | 90.4% | 94.6% | 99.5%  | 89.0% | 100.0% |
| 500+                           | 95.8% | 98.0% | 80.2% | 86.8% | 72.7% | 99.3%  | 83.4% | 93.3% | 99.4%  | 76.4% | 100.0% |

**Percentage of Discharges with a Verified Person Number (visitLink), All Available HCUP Databases within a State, 2004-2005**

| Type of Discharge              | AZ    | AR    | CA    | FL    | NC    | NE     | NV    | NY    | SC    | UT    | WA     |
|--------------------------------|-------|-------|-------|-------|-------|--------|-------|-------|-------|-------|--------|
| <b>Overall:</b>                | 97.3% | 96.4% | 89.5% | 95.0% | 63.2% | 99.4%  | 89.9% | 98.0% | 96.8% | 81.7% | 100.0% |
| <b>Age Group:</b>              |       |       |       |       |       |        |       |       |       |       |        |
| 0                              | 81.2% | 82.4% | 9.2%  | 13.6% | 17.3% | 100.0% | 54.0% | 91.1% | 97.0% | 44.1% | 100.0% |
| 1-17                           | 90.8% | 84.6% | 70.8% | 82.9% | 53.9% | 99.9%  | 66.6% | 94.9% | 97.4% | 41.6% | 100.0% |
| 18-44                          | 96.5% | 97.3% | 88.8% | 96.3% | 66.8% | 99.3%  | 95.2% | 96.8% | 96.5% | 92.7% | 100.0% |
| 45-64                          | 99.0% | 99.4% | 97.3% | 98.7% | 67.0% | 99.1%  | 98.2% | 97.4% | 96.9% | 95.7% | 100.0% |
| 65+                            | 99.4% | 99.7% | 98.6% | 99.2% | 67.7% | 99.5%  | 98.7% | 99.5% | 96.8% | 96.3% | 100.0% |
| <b>Expected Primary Payer:</b> |       |       |       |       |       |        |       |       |       |       |        |
| Medicare                       | 99.8% | 99.7% | 99.5% | 99.6% | 67.8% | 99.3%  | 99.2% | 99.0% | 96.7% | 96.1% | 100.0% |
| Medicaid                       | 92.5% | 90.4% | 67.4% | 81.6% | 53.3% | 99.4%  | 82.6% | 93.3% | 96.7% | 64.5% | 100.0% |
| Private insurance              | 95.7% | 96.0% | 79.3% | 92.5% | 64.6% | 99.5%  | 89.0% | 98.4% | 97.2% | 82.5% | 100.0% |
| Self-pay                       | 74.1% | 93.2% | 78.9% | 84.7% | 62.2% | 99.9%  | 80.5% | 96.0% | 95.6% | 82.6% | 100.0% |
| No charge                      | 74.3% | 92.1% | NA    | 89.2% | NA    | NA     | 83.0% | 87.1% | NA    | NA    | 100.0% |
| Other                          | 94.0% | 97.5% | 88.1% | 94.6% | 51.4% | 99.8%  | 88.3% | 97.9% | 98.7% | 83.7% | 100.0% |
| <b>Hospital Location:</b>      |       |       |       |       |       |        |       |       |       |       |        |
| Large central metropolitan     | 95.9% | NA    | 82.2% | 91.1% | 87.6% | NA     | 90.4% | 96.3% | NA    | 82.5% | 100.0% |
| Large fringe metropolitan      | 93.6% | 98.9% | 85.6% | 92.7% | 50.2% | NA     | NA    | 97.4% | 63.2% | 35.7% | 100.0% |
| Medium metropolitan            | 95.1% | 96.1% | 83.5% | 93.9% | 72.8% | 99.3%  | 88.9% | 98.7% | 99.0% | 83.4% | 100.0% |
| Small metropolitan             | 93.4% | 98.0% | 86.8% | 94.6% | 44.0% | NA     | NA    | 99.0% | 97.3% | 85.7% | 100.0% |
| Micropolitan                   | 95.7% | 96.5% | 90.9% | 97.4% | 66.1% | 99.7%  | 88.8% | 99.4% | 94.4% | 79.9% | 100.0% |
| Noncore                        | 89.3% | 95.4% | 92.3% | 98.5% | 28.9% | 99.6%  | 84.8% | 98.6% | 99.3% | 73.4% | 100.0% |
| <b>Hospital Bed Size:</b>      |       |       |       |       |       |        |       |       |       |       |        |
| <100                           | 95.4% | 96.1% | 86.2% | 96.5% | 54.2% | 99.7%  | 92.2% | 99.1% | 96.4% | 77.4% | 100.0% |
| 100-299                        | 94.1% | 95.6% | 82.6% | 93.9% | 67.2% | 99.7%  | 90.7% | 97.5% | 94.6% | 81.7% | 100.0% |
| 300-499                        | 97.7% | 98.0% | 83.1% | 93.5% | 70.8% | 98.7%  | 91.7% | 97.5% | 99.5% | 89.5% | 100.0% |
| 500+                           | 95.8% | 98.1% | 83.0% | 89.3% | 69.2% | 98.9%  | 87.9% | 96.4% | 99.6% | NA    | 100.0% |

**Percentage of Discharges with a Verified Person Number (visitLink), All Available HCUP Databases within a State, 2003-2004**

| Type of Discharge              | AZ    | CA    | NC    | NE     | NV    | NY    | SC    | UT    | WA     |
|--------------------------------|-------|-------|-------|--------|-------|-------|-------|-------|--------|
| <b>Overall:</b>                | 90.7% | 78.3% | 61.0% | 99.4%  | 90.5% | 97.7% | 96.8% | 82.0% | 100.0% |
| <b>Age Group:</b>              |       |       |       |        |       |       |       |       |        |
| 0                              | 62.0% | 6.4%  | 16.5% | 100.0% | 55.0% | 94.5% | 97.1% | 44.9% | 100.0% |
| 1-17                           | 74.8% | 65.0% | 51.5% | 99.9%  | 71.2% | 97.6% | 97.6% | 42.0% | 100.0% |
| 18-44                          | 93.7% | 85.4% | 63.7% | 99.2%  | 95.7% | 96.7% | 96.5% | 93.6% | 100.0% |
| 45-64                          | 98.4% | 96.6% | 65.1% | 99.2%  | 98.6% | 97.4% | 96.7% | 96.4% | 100.0% |
| 65+                            | 99.1% | 98.3% | 65.7% | 99.6%  | 99.0% | 99.6% | 96.7% | 96.8% | 100.0% |
| <b>Expected Primary Payer:</b> |       |       |       |        |       |       |       |       |        |
| Medicare                       | 99.7% | 99.5% | 65.6% | 99.3%  | 99.1% | 99.0% | 96.6% | 96.8% | 100.0% |
| Medicaid                       | 80.5% | 58.1% | 50.5% | 99.3%  | 85.8% | 93.6% | 96.9% | 65.8% | 100.0% |
| Private insurance              | 91.2% | 74.0% | 62.7% | 99.5%  | 89.1% | 98.8% | 97.2% | 82.3% | 100.0% |
| Self-pay                       | 64.2% | 71.0% | 55.2% | 99.7%  | 78.8% | 96.6% | 95.6% | 81.9% | 100.0% |
| No charge                      | 75.6% | NA    | NA    | NA     | 89.7% | 83.2% | NA    | NA    | 100.0% |
| Other                          | 89.7% | 83.5% | 55.5% | 99.9%  | 87.3% | 97.1% | 98.5% | 86.4% | 100.0% |
| <b>Hospital Location:</b>      |       |       |       |        |       |       |       |       |        |
| Large central metropolitan     | 91.5% | 77.5% | 88.9% | NA     | 91.1% | 96.8% | NA    | 82.4% | 100.0% |
| Large fringe metropolitan      | 90.1% | 81.1% | 50.5% | NA     | NA    | 98.0% | 61.4% | 58.3% | 100.0% |
| Medium metropolitan            | 88.8% | 78.0% | 70.0% | 99.4%  | 88.9% | 99.0% | 99.0% | 83.3% | 100.0% |
| Small metropolitan             | 87.2% | 81.3% | 42.1% | NA     | 89.4% | 99.3% | 97.3% | 85.3% | 100.0% |
| Micropolitan                   | 92.5% | 86.6% | 61.7% | 99.6%  | 85.4% | 99.6% | 94.6% | 78.6% | 100.0% |
| Noncore                        | 83.2% | 88.9% | 27.3% | 99.5%  | 87.7% | 99.0% | 99.3% | 77.2% | 100.0% |
| <b>Hospital Bed Size:</b>      |       |       |       |        |       |       |       |       |        |
| <100                           | 92.4% | 80.2% | 54.2% | 99.6%  | 91.5% | 99.4% | 96.2% | 78.2% | 100.0% |
| 100-299                        | 88.5% | 78.0% | 64.3% | 99.5%  | 90.7% | 98.1% | 94.7% | 81.6% | 100.0% |
| 300-499                        | 96.0% | 78.5% | 67.5% | 99.4%  | 91.2% | 98.1% | 99.6% | 89.7% | 100.0% |
| 500+                           | 89.2% | 77.7% | 66.9% | 98.8%  | 89.8% | 96.8% | 99.5% | NA    | 100.0% |

## APPENDIX E: SAS CODE FOR USAGE EXAMPLES

### Usage Example #1: Assigning Patient Characteristics

```
** Assigning Attributes

** combined event data -- combine all available data types for the years;
** for simplicity, the example uses only 2 attributes: AGE and ZIPINC_QRTL;
data NE_CombinedEvents1;
  keep KEY visitLink daysToEvent AGE ZIPINC_QRTL
       events sidEvents seddEvents sidSeddEvents sasdEvents
       verified sidVerified seddVerified sidSeddVerified sasdVerified;
merge NE_2005_DaysToEvent (in=_inDaysToEvent)
      NE_2006_DaysToEvent (in=_inDaysToEvent)
      NE_SID_2005_Core (in=_inSID_2005)
      NE_SID_2006_Core (in=_inSID_2006)
      NE_SASD_2005_Core (in=_inSASD_2005)
      NE_SASD_2006_Core (in=_inSASD_2006)
      NE_SEDD_2005_Core (in=_inSEDD_2005)
      NE_SEDD_2006_Core (in=_inSEDD_2006)
      end=lastObs;
by KEY;

** indicator for all events;
events = 1;
if _inDaysToEvent then verified = 1;
else verified = 0;

** indicator for SID events;
if _inSID then do;
  sidEvents = 1;
  if _inDaysToEvent then sidVerified = 1;
  else sidVerified = 0;
end;

** indicator for SEDD events;
if _inSEDD then do;
  seddEvents = 1;
  if _inDaysToEvent then seddVerified = 1;
  else seddVerified = 0;
end;

** indicator for SID/SEDD events;
sidSeddEvents = max(sidEvents, seddEvents);
sidSeddVerified = max(sidVerified, seddVerified);

** indicator for SASD events;
if _inSASD then do;
  sasdEvents = 1;
  if _inDaysToEvent then sasdVerified = 1;
  else sasdVerified = 0;
end;
run;

** sort the combined events into link order;
proc sort data=NE_CombinedEvents1;
  by visitLink daysToEvent;
run;
```

```

** use the first non-missing value for link ID attributes;
data NE_LinkAttribs2;
  keep visitLink foundAGE foundZIP;
  set NE_CombinedEvents1;
  by visitLink daysToEvent;
  where visitLink;  ** only process events with a visitLink ID;

  ** set attributes to missing for each new verified patient;
  length foundZIP $5;
  retain foundAGE foundZIP;
  if first.visitLink then do;
    foundAGE = .;
    foundZIP = '';
  end;

  ** select the first non-missing attribute;
  if foundAGE le .Z and AGE gt .Z then foundAGE = AGE;
  if foundZIP eq '' and ZIPINC_QRTL ne '' then foundZIP = ZIPINC_QRTL;

  ** create one set of attributes for each verified patient;
  if last.visitLink then output;
run;

** apply the consistent attributes to the event data;
data NE_CombinedEvents2;
  keep KEY visitLink daysToEvent AGE ZIPINC_QRTL
        events sidEvents seddEvents sidSeddEvents sasdEvents
        verified sidVerified seddVerified sidSeddVerified sasdVerified;
  merge NE_CombinedEvents1
        NE_LinkAttribs2 (in=_inLinkAttribs)
  by visitLink;

  ** apply uniform attributes to verified patient events;
  if _inLinkAttribs then do;
    AGE = foundAGE;
    ZIPINC_QRTL = foundZIP;
  end;
run;

```

## Usage Example #2: Revisits for Selected Patients

```
** Example 2, Follow-up Care
** Program code

** diagnosis CCS code to examine and clean period (months);
%let condX = 50;          ** diabetes mellitus with complications;
%let cleanPeriod = 6;    ** number of months for "clean" period with no condX claims;
%let firstYear = 2005;   ** first year of data;
%let lastYear = 2006;    ** last year of data;

** (1) combine event and revisit data -- keep only events with the specified condition;
data NE_EventsCondX1;
  keep KEY visitLink daysToEvent servSetting YEAR DQTR AMONTH LOS;
  merge NE_2005_DaysToEvent (in=_inDaysToEvent)
        NE_2006_DaysToEvent (in=_inDaysToEvent)
        NE_SID_2005_Core (in=_inSID)
        NE_SID_2006_Core (in=_inSID)
        NE_SASD_2005_Core (in=_inSASD)
        NE_SASD_2006_Core (in=_inSASD)
        NE_SEDD_2005_Core (in=_inSEDD)
        NE_SEDD_2006_Core (in=_inSEDD)
        end=lastObs;
  by KEY;

  ** (1.a) limit data to linkable patients;
  if _inDaysToEvent and (_inSID or _inSASD or _inSEDD);

  ** service type indicator;
  select;
    when (_inSID) servSetting = 'IP';
    when (_inSEDD) servSetting = 'ED';
    when (_inSASD) servSetting = 'AS';
    otherwise;
  end;

  ** (1.b) grab all events with condX;
  if _n_ eq 1 then put "Searched CCS diagnoses codes for values of '&condX'";
  array DXCCS {*} DXCCS1-DXCCS15;
  do i = 1 to NDX;
    if DXCCS{i} eq &condX then output NE_EventsCondX1;
  end;
run;

** (2) sort the combined events into visitLink (patient), service sequence order;
proc sort data=NE_EventsCondX1;
  by visitLink daysToEvent;
run;
```

```

** (3) find people with 2+ condX related events,
** calculate days between the 1st and 2nd event;
data NE_EventsCondX2;
  keep visitLink patientCount patientEvents servSetting1 servSetting2 days_EltoE2;
  set NE_EventsCondX1
    end=lastObs;
  by visitLink daysToEvent;

  retain patientCount 1;
  label patientCount = 'count of patients'
    patientEvents = 'number of events for patient'
    servSetting1 = 'type of service for the initial event'
    servSetting2 = 'type of service for the second event'
    days_EltoE2 = 'days between initial and second event';

  retain _condX_events;
  _condX_events + 1;
  _dayLag = lag(daysToEvent);

  retain _patients _pat_w2Plus _pat_wCleanPeriod _cleanPeriod days_EltoE2 patientEvents 0
    servSetting1 servSetting2;
  if first.visitLink then do;
    _patients + 1; ** count patients with condX;

    ** reset patient indicators and counters;
    days_EltoE2 = .;
    patientEvents = .;
    servSetting1 = ' ';
    servSetting2 = ' ';

    if not (first.visitLink and last.visitLink) then _pat_w2Plus + 1;

    ** (3.a.i) first claim must be after "cleanPeriod" month of first data year;
    if (YEAR gt &firstYear) or /* assumes clean period (months) < 12 */
      ( YEAR eq &firstYear and
        DQTR gt &cleanPeriod/3 and
        AMONTH gt &cleanPeriod and
        LOS lt &cleanPeriod*30 ) then do;
      _cleanPeriod = 1; ** indicator that patient had clean period;
      servSetting1 = servSetting;
      patientEvents = 1; ** counter for number of patient events;
    end; /* end-if (clean period) */
    else _cleanPeriod = 0;
  end; /* end-if (first visitLink) */

  else do; /* not first visitLink */
    ** (3.b) revisit - clean period already found;
    if _cleanPeriod then do;
      patientEvents + 1; ** count number of events for this patient;
      if patientEvents eq 2 then do;
        days_EltoE2 = daysToEvent - _dayLag;
        servSetting2 = servSetting;
        _pat_wCleanPeriod + 1; ** count patients with a clean period;
      end; /* end-if (second event) */
    end; /* end-if (clean period) */

    else do; /* clean not (yet) period found */
      ** (3.a.ii) no clean pd yet found - check lag days f/ clean pd between events;
      if _dayLag/30 gt &cleanPeriod then do; /* 99% correct */
        _cleanPeriod = 1; ** indicator that patient had clean period;
        servSetting1 = servSetting;
        patientEvents = 1; ** counter for number of patient events;
      end; ** end-if (lag clean period);
    end; /* end-else (clean period not found) */
  end; /* end-else (not first visitLink) */

  ** output one obs per visitLink;
  if last.visitLink and patientEvents ge 2 then output;

```



```

** (3.c) summarize processing;
if lastObs then do;
  put '=====';
  put "Processing summary -- events with diagnosis (CCS) category: &condX";
  put '-started with ' _condX_events 'events (total)';
  put '      for ' _patients 'patients.';
  put '-there were ' _pat_w2Plus "patients with 2+ DXCCS '&condX' events";
  put "-clean period, no claim with DXCCS '&condX' for at least &cleanPeriod months";
  put ' before the first claim: ' _pat_wCleanPeriod ;
  put '=====';
  put /;
end;
run;

proc format;
  picture pctfmt low-high='009 %';
run;

** (4) statistics for the number of condX related events;
title1 "Number of DXCCS &condX Events - Distribution";
title2 "for Patients with Multiple DXCCS &condX Events";
proc means data=NE_EventsCondX2 maxdec=2 mean p25 p50 p75 max;
  var patientEvents;
run;

** (5) summarize revisits by the initial and second service settings;
title "Count and Days for DXCCS '&condX' Patients with Multiple Hospital Events";
proc tabulate data=NE_EventsCondX2 format=comma12.;
  class servSetting1 servSetting2 /descending;
  var patientCount;
  table (servSetting1 all),
        (servSetting2 all)*
        (patientCount*sum*f=comma12.
         days_EltoE2*(mean median)*f=8.1);
run;

```

# Output

## Number of DXCCS 50 Events - Distribution for Patients with Multiple DXCCS 50 Events

The MEANS Procedure

Analysis Variable : patientEvents number of events for patient

| Mean | 25th Pctl | 50th Pctl | 75th Pctl | Maximum |
|------|-----------|-----------|-----------|---------|
| 4.14 | 2.00      | 3.00      | 4.00      | 97.00   |

## Count and Days for DXCCS '50' Patients with Multiple Hospital Events

|               | type of service for the second event |                                       |        |                   |                                       |        |                   |                                       |        |                   |                                       |        |
|---------------|--------------------------------------|---------------------------------------|--------|-------------------|---------------------------------------|--------|-------------------|---------------------------------------|--------|-------------------|---------------------------------------|--------|
|               | IP                                   |                                       |        | ED                |                                       |        | AS                |                                       |        | All               |                                       |        |
|               | count of patients                    | days between initial and second event |        | count of patients | days between initial and second event |        | count of patients | days between initial and second event |        | count of patients | days between initial and second event |        |
|               | Sum                                  | Mean                                  | Median | Sum               | Mean                                  | Median | Sum               | Mean                                  | Median | Sum               | Mean                                  | Median |
| initial event |                                      |                                       |        |                   |                                       |        |                   |                                       |        |                   |                                       |        |
| IP            | 1,850                                | 61.6                                  | 6.0    | 367               | 102.8                                 | 35.0   | 157               | 140.3                                 | 111.0  | 2,374             | 73.2                                  | 14.0   |
| ED            | 254                                  | 122.0                                 | 75.0   | 299               | 93.3                                  | 34.0   | 40                | 123.0                                 | 86.5   | 593               | 107.6                                 | 52.0   |
| AS            | 158                                  | 151.3                                 | 120.0  | 24                | 114.0                                 | 115.5  | 466               | 41.6                                  | 14.0   | 648               | 71.0                                  | 21.0   |
| All           | 2,262                                | 74.7                                  | 16.0   | 690               | 99.1                                  | 37.0   | 663               | 69.9                                  | 21.0   | 3,615             | 78.5                                  | 21.0   |

### Usage Example #3: Preceding Visits to any Hospital Setting for Selected Patients

```
** Example 3, Preceding Events
** Program code

** procedure CCS code to examine;
%let servX = 44;          ** CABG;

** combine event and revisit data -- flag events with the specified service;
data TN_AllEvents;
  keep KEY visitLink daysToEvent servX dxCcs1 servSetting events;
  merge TN_2005_DaysToEvent (in=_inDaysToEvent)
        TN_2006_DaysToEvent (in=_inDaysToEvent)
        TN_SID_2005_Core (in=_inSID)
        TN_SID_2006_Core (in=_inSID)
        TN_SASD_2005_Core (in=_inSASD)
        TN_SASD_2006_Core (in=_inSASD)
        TN_SEDD_2005_Core (in=_inSEDD)
        TN_SEDD_2006_Core (in=_inSEDD)
        end=lastObs;
  by KEY;
  if _inSID or _inSASD or _inSEDD;

  ** flag events with the specified service;
  array prCcs {*} prCcs1-prCcs30;
  do i = 1 to dim(prCcs);
    if prCcs{i} eq &servX then do;
      servX = 1;
      leave;
    end;
  end;

  array cptCcs {*} cptCcs1-cptCcs30;
  do i = 1 to dim(cptCcs);
    if cptCcs{i} eq &servX then do;
      servX = 1;
      leave;
    end;
  end;

  ** service type indicator;
  select;
    when (_inSID) servSetting = 'IP';
    when (_inSEDD) servSetting = 'ED';
    when (_inSASD) servSetting = 'AS';
    otherwise;
  end;

  ** counters;
  retain events 1;
  if _inSID then sidEvents = 1;
  if _inSEDD then seddEvents = 1;
  if _inSASD then sasdEvents = 1;
run;

** sort the combined events into visitLink (patient), service sequence order;
proc sort data=TN_AllEvents out=TN_OrderedEvents (index=(servX));
  by visitLink daysToEvent;
run;
```

```

** find people who received the topic service;
proc sql;
  create table TN_TopicPop as
    select visitLink,
           min(daysToEvent) as firstService
    from TN_OrderedEvents
    where servX eq 1
    group by visitLink;
quit;

** for people with the service, select all events prior to the service;
data TN_PriorEvents;
  merge TN_OrderedEvents (in=_inEventData)
        TN_TopicPop (in=_inTopicPop);
  by visitLink;
  if _inTopicPop;
run;

** summarize prior events by primary diagnosis and setting;
title 'Count of patients by Primary Diagnosis and Service Setting';
proc tabulate data=TN_PriorEvents format=commal2.;
  class dxCcs1 servSetting;
  format DXCCS1 FDCCSPDX.;
  var events;
  table (all DXCCS1),
        (all servSetting)*(events*sum);
run;

```

**Output**

**Count of patients by Primary Diagnosis and Service Setting**

|  | servSetting |        |         |         |
|--|-------------|--------|---------|---------|
|  | All         | AS     | ED      | IP      |
|  | events      | events | events  | events  |
|  | Sum         | Sum    | Sum     | Sum     |
|  |             |        |         |         |
| All                                      | 613,541     | 87,348 | 292,336 | 233,857 |
| CCS: principal diagnosis                 |             |        |         |         |
| 1: Infectious and Parasitic DX           | 8,351       | 273    | 6,726   | 1,352   |
| 2: Neoplasms                             | 27,723      | 24,110 | 221     | 3,392   |
| 3: Endocr, Nutri, Metab, Immun DX        | 7,074       | 1,626  | 3,095   | 2,353   |
| 4: Dx of Blood, Blood-Forming Organs     | 2,689       | 416    | 1,510   | 763     |
| 5: Mental Disorders                      | 8,004       | 216    | 5,939   | 1,849   |
| 6: Dx of Nervous System, Sense Organs    | 46,362      | 9,754  | 34,891  | 1,717   |
| 7: Dx of Circulatory System              | 72,483      | 13,626 | 18,290  | 40,567  |
| 8: Dx of Respiratory System              | 59,376      | 7,848  | 43,290  | 8,238   |
| 9: Dx of Digestive System                | 30,309      | 5,802  | 18,551  | 5,956   |
| 10: Dx of Genitourinary System           | 17,253      | 3,260  | 11,766  | 2,227   |
| 11: Complic Preg, Birth, Puerperium      | 15,170      | 2,244  | 4,302   | 8,624   |
| 12: Dx of Skin and Subcutaneous Tissue   | 8,553       | 826    | 6,626   | 1,101   |
| 13: Dx of Musculoskel, Connective Tissue | 25,937      | 6,453  | 16,583  | 2,901   |
| 14: Congenital Anomalies                 | 2,098       | 917    | 180     | 1,001   |
| 15: Perinatal Conditions                 | 145,485     | 810    | 3,227   | 141,448 |
| 16: Injury and Poisoning                 | 90,679      | 3,639  | 79,694  | 7,346   |
| 17: Other Conditions                     | 45,995      | 5,528  | 37,445  | 3,022   |